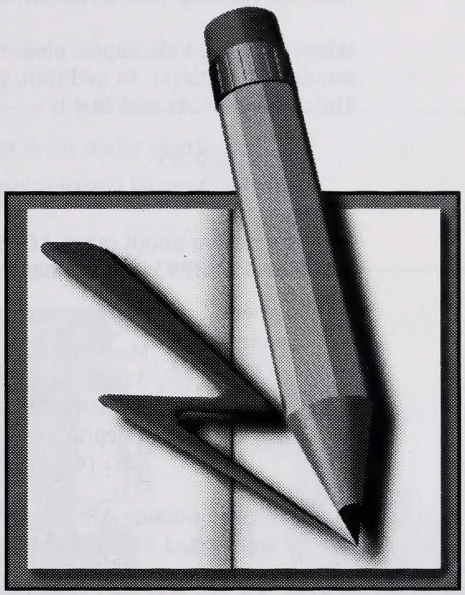
1997-403 c2 v1
University of Alberta Library
0 1620 3532755 8

MATH FACTOR

JAN 20 1997

STATISTICS



Lesson 1:
Normal Distribution

Student Workbook

Produced by

THE EDUCATION STATION ACCESS

for

Alberta EDUCATION

This guide accompanies **Normal Distribution** from the Math Factor series.

Writer: Jerry Polei
Editor: Wilma Gonsalves
Producer: Katherine Rankin

Educators in public educational institutions in Alberta may verbally quote, or use in exams, excerpts from this publication, except material that has been credited to another source. Additional copies of the guide can be obtained, at cost plus handling and postage, from the ACCESS Media Resource Centre.

When ordering videotapes, please use the Basic Program Number (BPN), plus the program number and title(s). In addition, please attach a purchase order to cover duplication fees. The order number and title is

BPN #3689-45
Statistics: Normal Distribution

For information about other ACCESS materials, the toll-free number for Alberta educators is 1-800-352-8293, or you may write to the following address:

ACCESS
Media Resource Centre
3720 - 76 Avenue
EDMONTON, Alberta T6B 2N9
Telephone: (403) 440-7729
Fax: (403) 440-8899

Inquiries from outside Alberta regarding purchase of this series or other ACCESS materials should be directed to ACCESS Program Sales at the address given above.

© 1996 Alberta Education

ISBN 1-896762-42-5

PRINTED AND BOUND IN CANADA

UNIVERSITY LIBRARY
UNIVERSITY OF ALBERTA

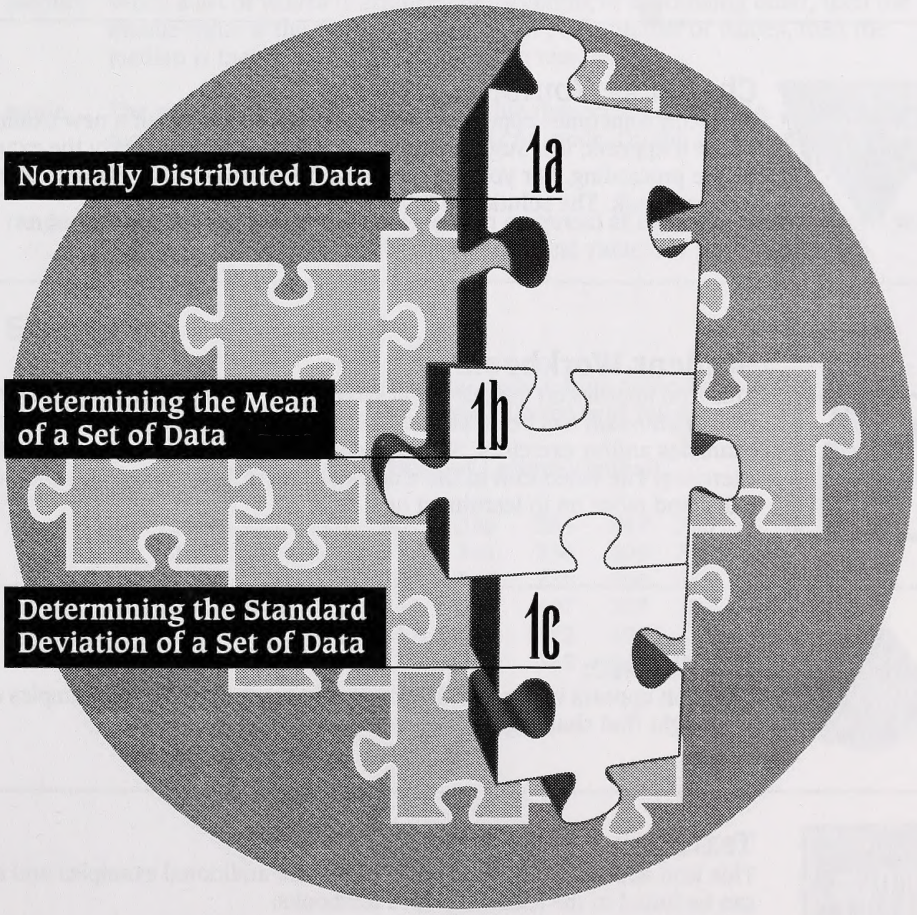
Lesson 1: NORMAL DISTRIBUTION

LEARNER EXPECTATIONS

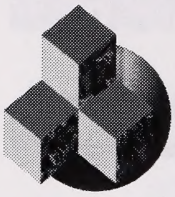
By the end of Lesson 1 you should be able to

- identify whether or not a set of data is normally distributed
- find and interpret the mean and standard deviation of a set of normally distributed data

This lesson has three parts:



EXPLANATION OF ICONS



Building Blocks:

Do you remember most of your Mathematics 20? This course builds upon many of those skills and concepts. The Building Blocks icon appears in the workbook to indicate review examples and/or exercises. To refresh your memory, do this section in the workbook before watching the video.

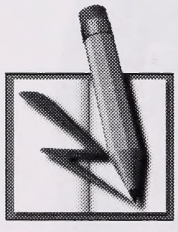
Video:

Whenever this icon appears in the workbook, you should watch the indicated segment of the video program. Preceding this icon, there is often a problem titled "Challenge yourself." Try the problem before watching the video. As you work through a program segment, stop the video to

- challenge yourself to find the solution before it is given
- write down key ideas and examples
- reflect on what is being taught

Challenge Yourself:

This icon sometimes appears on the television screen when a new example is starting. When it appears, it is suggested that you stop the video and try the example yourself before proceeding. For your convenience, sometimes the example is already written in the workbook. The solution will be given in the video.



Student Workbook:

This icon appears on the television screen at the end of teaching segments, when a new skill or sub-skill has been taught. The icon indicates that the workbook contains related examples and/or exercises. Stop the video to study the examples and complete the exercises. The video icon in the workbook will indicate when you should return to the video and move on to learning a new skill.

Excellence:

This icon appears in the workbook to indicate advanced level examples or exercises that you might find challenging. Test your new skills!

Textbook Reference:

This icon appears in the workbook to indicate additional examples and exercises that can be found in the following three textbooks:

- *Mathematics 12 (Addison-Wesley)*
- *Mathematics: Principles and Process 12 (Nelson Canada)*
- *Holtmath 12 (Holt, Rinehart and Winston of Canada, Limited)*

Building Blocks

Do you remember how to organize a set of data using a stem-and-leaf plot or a frequency distribution table? Can you use the information in a frequency distribution table to draw a histogram, a frequency polygon or a frequency curve? Before watching this program, refresh your memory by studying the example and doing the exercises.

The branch of mathematics called **statistics** deals with the **collection** of information, called **data**, the **organization**, **presentation** and the **interpretation** of the **data**.

mean The mean, or what is commonly referred to as the average, is found by adding all data values and then dividing this sum by the total number of data values.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} \text{ or } \bar{x} = \frac{\sum x}{n} \text{ or } \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

median When a set of values is arranged in ascending or descending order, then the middle value is the median. If there is an even number of values, then the median is the average of the middle two numbers.

mode The mode of a set of values is the value that occurs most often. If each value of the set occurs only once, then there is no mode. On the other hand, there may be more than one mode in a set.

range The range is a measure of dispersion or spread of the data. The range of a set of values is the difference between the smallest value and the largest value.

Example

Forty-eight heads of lettuce were selected at random from a one–hectare plot, and the mass of each head was determined. The data collected is as follows:

Mass of Each Head of Lettuce (grams)

346	325	295	402	386	354	347	333
389	352	409	372	361	356	409	398
409	355	376	382	345	289	366	349
327	387	356	349	368	407	287	354
348	359	348	362	337	372	392	315
367	362	373	359	352	332	354	361

Complete the following to organize, present and interpret the data.

- Use a stem-and-leaf plot and determine the measures of central tendency (the mean, median and mode) and the range.
- Use a frequency distribution table.
- Construct a histogram, a frequency polygon and a frequency curve.

Solution

- The stem-and-leaf plot is as follows:

Mass of Each Head of Lettuce (grams)

Stem	Leaf
28	7, 9
29	5
30	
31	5
32	5, 7
33	2, 3, 7
34	5, 6, 7, 8, 8, 9, 9
35	2, 2, 4, 4, 4, 5, 6, 6, 9, 9
36	1, 1, 2, 2, 6, 7, 8
37	2, 2, 3, 6
38	2, 6, 7, 9
39	2, 8
40	2, 7, 9, 9, 9

Note: The data in each leaf is arranged in ascending order. This makes it easier to determine measures of central tendency and the range.

The mean is $\frac{17\,233}{48} \doteq 359$.

The median is $\frac{(356 + 359)}{2} = 357.5$.

For this example, the data is said to be bimodal; that is, there are two modes. The modes are 354 and 409 because each of these values appears three times in the data.

The range for this example is $409 - 287 = 122$.

- Represent the data using a frequency distribution table.

Step 1: Find the range of the data.

$$\text{Range} = 409 - 287 = 122$$

Step 2: Decide on the number of classes and find the class width. Suppose it was decided that seven classes would be sufficient.

Class width is equal to the range divided by the number of classes.

$$\frac{122}{7} \doteq 17.4$$

$$= 18 \quad \text{(rounded to the next highest whole number)}$$

The lower limit of the first class would be 287, the lowest data value.

The upper limit of the first class would be 304, the lowest data value increased by one less than the class width.

The second class would have a lower limit of 304 + 1 = 305 and an upper limit of 305 + 17 = 322.

Subsequent class limits are determined in a similar manner.

Step 3: Set up a table with eight rows and five columns. Name the table and columns as shown; then enter the class limits and each class midpoint. (To get the midpoint, add the two extremes of the class and divide by two.)

Step 4: Classify all data into one of the seven classes by using tally marks.

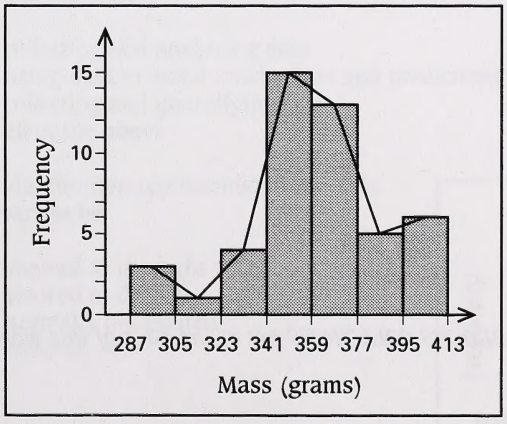
Step 5: Enter the frequency for each class. The total of the frequency column should equal the number of data values given.

Frequency Distribution Table for Lettuce Head Masses

Class	Class Limit	Class Midpoint	Tally	Frequency
1	287 to 304	295.5	///	3
2	305 to 322	313.5	/	1
3	323 to 340	331.5	𝍸	5
4	341 to 358	349.5	𝍸 𝍸 𝍸	15
5	359 to 376	367.5	𝍸 𝍸 ///	13
6	377 to 394	385.5	𝍸	5
7	395 to 412	403.5	𝍸 /	6

- Represent the data using a histogram, frequency polygon and a frequency curve.

Histogram and Frequency Polygon



The frequency curve is plotted in a manner similar to the frequency polygon, but a smooth curve is used.

Additional Examples
Mathematics 12: p. 433
Mathematics: Principles and Process 12: p. 434
Holtmath 12: p. 376

Exercises

1. Organize the data using a stem-and-leaf plot.

   A sample of 40 people measured their pulse for one minute. The data collected is shown.

Pulses of 40 People

66	61	79	71	53	86	81	73	84	87	76	64	67	83	71	68	81	92	86	56
62	67	79	76	82	93	78	62	76	76	59	82	96	72	77	88	76	90	73	71

2. Use five classes to organize the data in question 1 into a frequency distribution table; then represent the data using a histogram, a frequency polygon and a frequency curve.

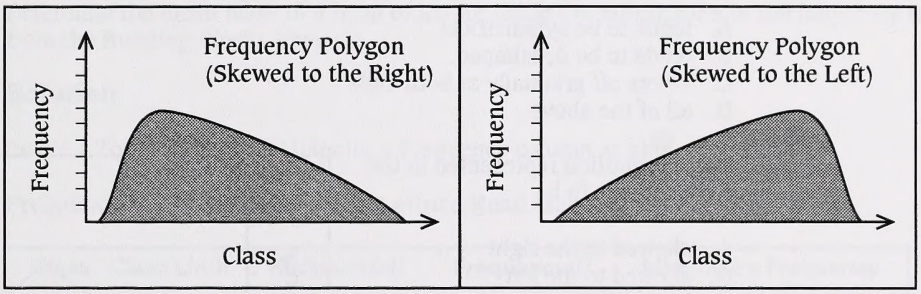
BUILDING BLOCKS SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics 12: p. 435 (Questions 5, 6, 8, 10)
Mathematics: Principles and Process 12: p. 435 (Question 4)
Holtmath 12: p. 377 (Questions 2-6)

Normally Distributed Data

Watch the Introduction and Video Segment 1a.

Skewness



A frequency polygon that is skewed to the right is characterized by a distribution that rises quickly and then comes back to the horizontal more slowly; that is, the "right tail" is longer than its "left tail."

A frequency polygon that is skewed to the left is characterized by a distribution that rises slowly and then drops quickly back to the horizontal. The "left tail" is longer than the "right tail."

Exercises

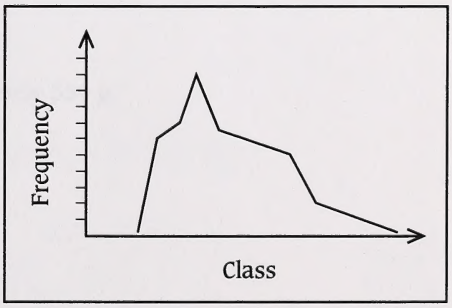
Select the best answer for each of the following:

1. Statistics involves

   A. collecting and analyzing data
   B. using data to make conclusions and predictions
   C. collecting and quantifying data
   D. all of the above

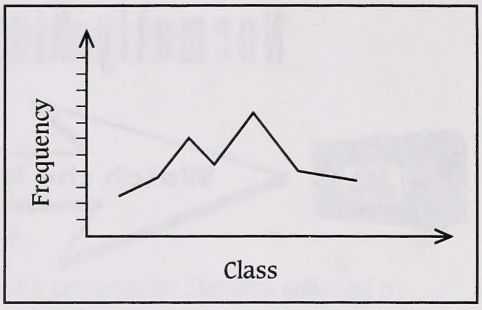
2. The distribution represented to the right appears to be

   A. skewed to the right
   B. skewed to the left
   C. normally distributed
   D. bimodal



3. The distribution represented to the right appears to be

   A. skewed to the right
   B. skewed to the left
   C. normally distributed
   D. bimodal

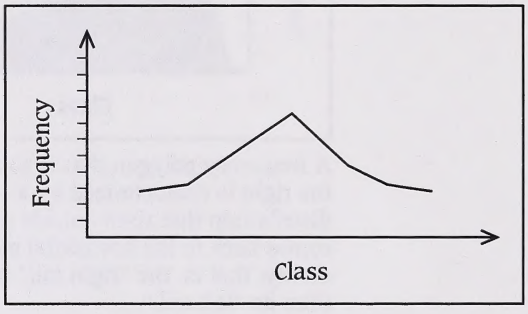


4. A distribution that can be modelled as a normal curve

   A. tends to be symmetrical
   B. tends to be bellshaped
   C. tapers off gradually at both ends
   D. all of the above

5. The distribution represented to the right appears to be

   A. skewed to the right
   B. skewed to the left
   C. normally distributed
   D. bimodal



EXERCISE 1a SOLUTIONS are at the back of this workbook.

Determining the Mean of a Set of Data

Watch Video Segment 1b.

Example

Determine the mean mass of a head of lettuce using the information in the frequency table from the Building Blocks Example.

Solution

Create a Totals row and a Midpoint × Frequency column as shown.

Frequency Distribution Table for Lettuce Head Masses

Class	Class Limit	Midpoint *(xi)*	Frequency *(f)*	Midpoint × Frequency
1	287 to 304	295.5	3	886.5
2	305 to 322	313.5	1	313.5
3	323 to 340	331.5	5	1657.5
4	341 to 358	349.5	15	5242.5
5	359 to 376	367.5	13	4777.5
6	377 to 394	385.5	5	1927.5
7	395 to 412	403.5	6	2421
Totals			48	17 226.0

The mean is then determined by dividing the total of the Midpoint × Frequency column by the total of the Frequency column.

$$\bar{x} = \frac{\left(\sum f \times x_i\right)}{n}$$
$$= \frac{17\ 226}{48}$$
$$\doteq 358.9$$

The mean mass of a head of lettuce is approximately 359 g.

Additional Examples
Mathematics 12: pp. 439, 440
Mathematics: Principles and Process 12: p. 436
Holtmath 12: pp. 388, 389

Exercises

6. The following chart represents the number of motorcycles sold on various days in the month of July at the Rev Up and Go dealership. Determine the mean number of motorcycles sold per day.

Class	Number Sold	Frequency
1	1 to 3	2
2	4 to 6	3
3	7 to 9	5
4	10 to 12	10
5	13 to 15	6
6	16 to 18	3
7	19 to 21	2

7. Determine the mean, median and mode for the data illustrated in the following stem-and-leaf plot.

Student Marks in a Math Exam (expressed as percents)

Stem	Leaf
5	3, 6, 9
6	6, 1, 4, 7, 8, 2, 7, 2
7	9, 1, 3, 6, 1, 9, 6, 8, 6, 6, 2, 7, 6, 3, 1
8	6, 1, 4, 7, 3, 1, 6, 2, 2, 8
9	2, 3, 6, 0

EXERCISE 1b SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics 12: p. 440 (Questions 1-4, 7a, 7b, 9a, 9b, 9c)
Mathematics: Principles and Process 12: p. 437 (Questions 1, 2, 4, 6)
Holtmath 12: p. 377 (Questions 5, 6), p. 381 (Questions 2, 3)

Determining the Standard Deviation of a Set of Data

challenge yourself

What does σ_{n-1} represent?

Watch Video Segment 1c.

The range and the standard deviation are both measures of dispersion. The range indicates the difference between the highest and lowest data values, but it does not tell much about how the data are distributed. The standard deviation tells how the data are distributed around the mean.

Example

Determine the standard deviation for the following sample data.

Lettuce Head Masses (g) 400 415 430 460 480

Solution

Step 1: Determine the mean.

$$\bar{x} = \frac{\left(\sum X\right)}{n}$$
$$= \frac{2185}{5}$$
$$= 437$$

Step 2: Complete the table.

Data Value (X)	Mean $(\bar{x})$	$(X - \bar{x})$	$(X - \bar{x})^2$
400	437	−37	1369
415	437	−18	324
430	437	−7	49
460	437	23	529
480	437	43	1849
Totals			4120

Sample Standard Deviation

$$\sigma_{n-1} = \sqrt{\frac{\sum(X-\bar{x})^2}{(n-1)}}$$
$$= \sqrt{\frac{4120}{(5-1)}}$$
$$\doteq 32.1 \text{ g}$$

Population Standard Deviation

If the data is a population, then

$$\sigma_n = \sqrt{\frac{\sum(X-\bar{x})^2}{n}} \quad \text{or} \quad \sigma_n = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)}{n}}$$

Note: If you are not specifically asked for the sample standard deviation, determine the population standard deviation.

Note: Refer to the instruction manual for your calculator to see how you can quickly calculate the mean and standard deviation. You may get a slightly different answer when you use a calculator because calculators round numbers in a different way.

Additional Examples
Mathematics 12: p. 446
Mathematics: Principles and Process 12: p. 441, 442
Holtmath 12: p. 395

Exercises

8. Fill in the blanks.

   a. From the video, the mean of a population would be referred to as a _____.

   b. From the video, the mean of a sample would be referred to as a _____.

   c. A formula to calculate the standard deviation of a sample is _____.

   d. A formula to calculate the standard deviation of a population is _____.

9. State whether the following are true (T) or false (F).

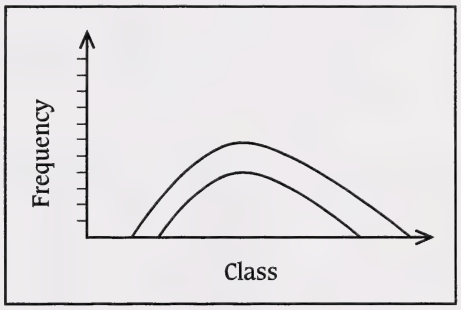
____ a. The standard deviation of a sample will equal the standard deviation of the population from which the sample was taken.

____ b. If two normally distributed populations have the same mean, then they will definitely have the same standard deviation.

____ c. The two distributions represented in the diagram to the right have means that are approximately equal.

____ d. The two distributions represented in the diagram to the right have standard deviations that are approximately equal.

____ e. The two distributions represented in the diagram to the right could be modelled using a standard normal curve.



10. Use the statistics mode of your calculator to determine the mean and standard deviation for the following set of final scores that a basketball team has had in 12 different games.

102, 97, 105, 105, 96, 111, 106, 100, 99, 103, 102, 98

EXERCISE 1c SOLUTIONS are at the back of this workbook.

Additional Exercises
Mathematics 12: p. 448 (Questions 6a, 6b, 6c, 6d, 6e, 7a, 7b, 7d, 12)
Mathematics: Principles and Process 12: p. 442 (Questions 4-6)
Holtmath 12: p. 395 (Questions 10, 12)

WHAT HAVE YOU LEARNED?
Return to the Learner Expectations at the beginning of this lesson and check off the points that you have mastered. You are barraged with statistics on a daily basis. Many decisions you make are based on these statistics. Understanding how the mean and standard deviation of a set of data are calculated will assist you in making wise decisions. As well, you need to have these skills as you continue on to **Lesson 2: Characteristics of Normal Distribution**.

Solutions

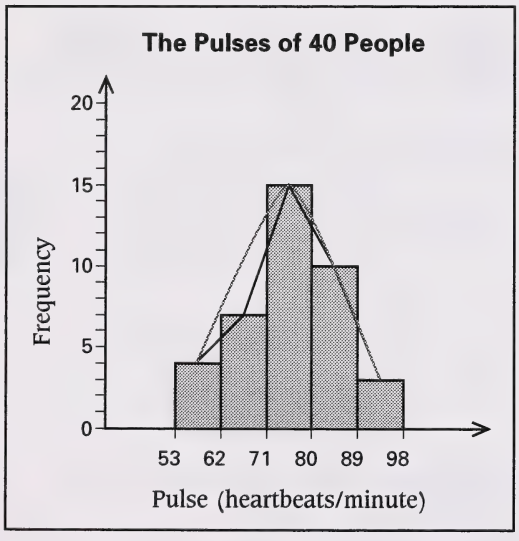
BUILDING BLOCKS

1.

Stem	Leaf
5	3, 6, 9
6	1, 2, 2, 4, 6, 7, 7, 8
7	1, 1, 1, 2, 3, 3, 6, 6, 6, 6, 6, 7, 8, 9, 9
8	1, 1, 2, 2, 3, 4, 6, 6, 7, 8
9	0, 2, 3, 6

2.

Class	Class Width	Midpoint	Tally	Frequency
1	53 to 61	57	////	4
2	62 to 70	66	𝍸 //	7
3	71 to 79	75	𝍸 𝍸 𝍸	15
4	80 to 88	84	𝍸 𝍸	10
5	89 to 98	93	////	4



EXERCISE 1a

1. D
2. A
3. D
4. D
5. C

EXERCISE 1b

6. The mean number of motorcycles sold per day is approximately 11.

7. The mean of the student marks is 75.5%; the median and mode of student marks are both 76%.

EXERCISE 1c

8. a. parameter

   b. statistic

   c. $\sigma_{n-1} = \sqrt{\dfrac{\sum(X-\bar{x})^2}{(n-1)}}$

   d. $\sigma_n = \sqrt{\dfrac{\sum(X-\bar{x})^2}{n}}$ or see the second formula on p. 12

9. a. F

   b. F

   c. T

   d F

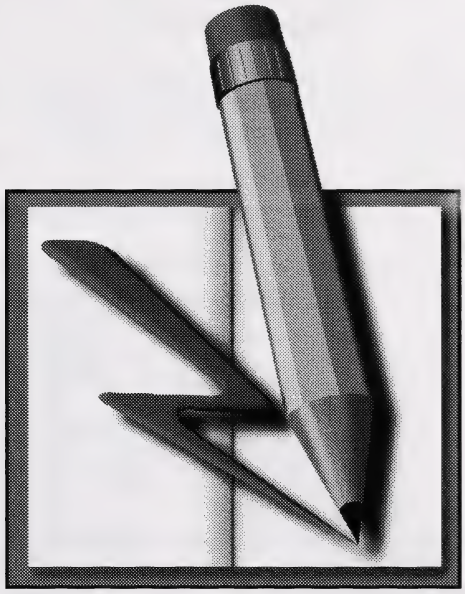
   e. T

10. The mean is 102, and the standard deviation is approximately 4.14.

2.1997-403cor v2

MATH FACTOR

STATISTICS



Lesson 2:
Characteristics of Normal Distribution

Student Workbook

Produced by

THE EDUCATION STATION ACCESS

for

Alberta EDUCATION

This guide accompanies **Characteristics of Normal Distribution** from the Math Factor series.

Writer: Jerry Polei
Editor: Wilma Gonsalves
Producer: Katherine Rankin

Educators in public educational institutions in Alberta may verbally quote, or use in exams, excerpts from this publication, except material that has been credited to another source. Additional copies of the guide can be obtained, at cost plus handling and postage, from the ACCESS Media Resource Centre.

When ordering videotapes, please use the Basic Program Number (BPN), plus the program number and title(s). In addition, please attach a purchase order to cover duplication fees. The order number and title is

BPN #3689-46
Statistics: Characteristics of Normal Distribution

For information about other ACCESS materials, the toll-free number for Alberta educators is 1-800-352-8293, or you may write to the following address:

ACCESS
Media Resource Centre
3720 - 76 Avenue
EDMONTON, Alberta T6B 2N9
Telephone: (403) 440-7729
Fax: (403) 440-8899

Inquiries from outside Alberta regarding purchase of this series or other ACCESS materials should be directed to ACCESS Program Sales at the address given above.

© 1996 Alberta Education

ISBN 1-896762-43-3

PRINTED AND BOUND IN CANADA

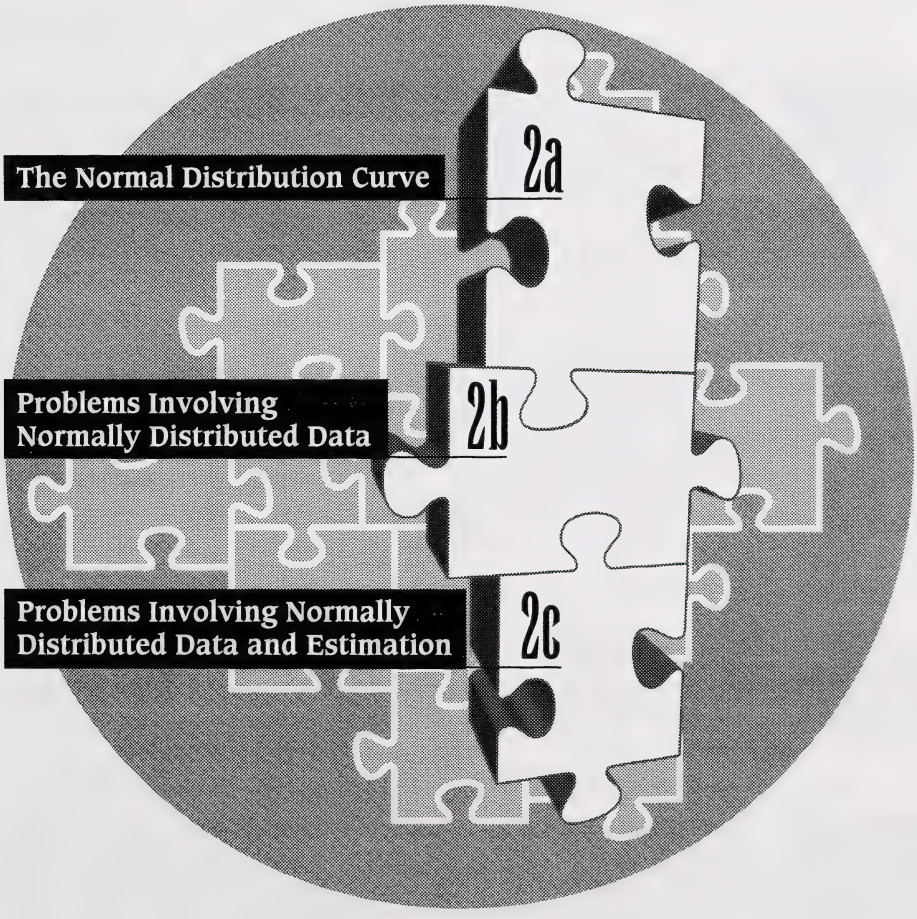
Lesson 2: CHARACTERISTICS OF NORMAL DISTRIBUTION

LEARNER EXPECTATIONS

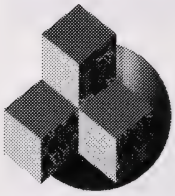
By the end of Lesson 2 you should be able to

- apply the characteristics of a normal distribution
- solve problems involving data that is normally distributed

This lesson has three parts:



EXPLANATION OF ICONS



Building Blocks:

Do you remember most of your Mathematics 20? This course builds upon many of those skills and concepts. The Building Blocks icon appears in the workbook to indicate review examples and/or exercises. To refresh your memory, do this section in the workbook before watching the video.

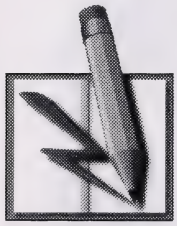
Video:

Whenever this icon appears in the workbook, you should watch the indicated segment of the video program. Preceding this icon, there is often a problem titled "Challenge yourself." Try the problem before watching the video. As you work through a program segment, stop the video to

- challenge yourself to find the solution before it is given
- write down key ideas and examples
- reflect on what is being taught

Challenge Yourself:

This icon sometimes appears on the television screen when a new example is starting. When it appears, it is suggested that you stop the video and try the example yourself before proceeding. For your convenience, sometimes the example is already written in the workbook. The solution will be given in the video.



Student Workbook:

This icon appears on the television screen at the end of teaching segments, when a new skill or sub-skill has been taught. The icon indicates that the workbook contains related examples and/or exercises. Stop the video to study the examples and complete the exercises. The video icon in the workbook will indicate when you should return to the video and move on to learning a new skill.

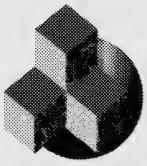
Excellence:

This icon appears in the workbook to indicate advanced level examples or exercises that you might find challenging. Test your new skills!

Textbook Reference:

This icon appears in the workbook to indicate additional examples and exercises that can be found in the following three textbooks:

- *Mathematics 12 (Addison-Wesley)*
- *Mathematics: Principles and Process 12 (Nelson Canada)*
- *Holtmath 12 (Holt, Rinehart and Winston of Canada, Limited)*



Building Blocks

Do you remember how to determine the probability of a single event or a compound event? Before watching this program, refresh your memory by studying the following examples and doing the exercises.

Probabilities can be determined in two ways:

- theoretically, where probabilities are assigned on the basis of the total number of outcomes in a sample space.
- experimentally, where a random sample of a population is selected and the probabilities of this random sample are assigned to the entire population

The methods discussed in the following examples can be used to solve both theoretical and experimental probability problems.

Example 1

If a die is tossed, determine the probability of the following:

- a six will come up
- a seven will come up
- a value less than seven will come up
- a six or a four will come up
- a value that is even or prime will come up

Solution

- a six will come up

  The probability of an event is the number of favorable outcomes divided by the total number of outcomes.

  There are six possible outcomes. A sample space or representation of all possible outcomes is $\{1, 2, 3, 4, 5, 6\}$. Let A be the event that a six comes up; then the probability of A, denoted $P(A)$, is $\frac{1}{6}$.

- a seven will come up

  This is an impossible event because a 7 does not exist on a die. The probability of an impossible event is zero.

  $P(\text{impossible event}) = 0$

  If A is the event that a seven comes up, then the $P(A)$ is 0.

- a value less than seven will come up

  Any value will be less than seven; therefore, this is a certain event. The probability of a certain event is one.

  $P(\text{certain event}) = 1$

  If A is an event of a sample space, then $0 \leq P(A) \leq 1$.

  If A is the event that a value less than seven will come up, then the $P(A)$ is 1.

- a six or a four will come up

  The probability of this event is $\frac{2}{6} = \frac{1}{3}$ (the two favorable outcomes {4, 6} divided by the total number of possible outcomes).

  This particular problem can also be solved using the theory of compound events. If A = {4} and B = {6}, then you are interested in the P(A or B), the probability of A or B. These two events are mutually exclusive; they cannot both happen at the same time.

  If two events are mutually exclusive, P(A or B) = P(A) + P(B).

  In this particular case, $P(A \text{ or } B) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$.

- a value that is even or prime will come up

  The probability of this event is $\frac{5}{6}$ (five favorable outcomes {2, 3, 4, 5, 6} divided by the total number of possible outcomes).

  Again, compound event theory can be used to solve the problem.

  Let A represent the event that the value which comes up is even, A = {2, 4, 6}. Let B represent the event that the value which comes up is prime, B = {2, 3, 5}. Clearly, the two events are not mutually exclusive. The number 2 is both even and prime. Both events can occur at the same time.

  When two events are not mutually exclusive, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

  For the question, $P(A \text{ or } B) = \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}$. (Subtract $\frac{1}{6}$ because one of the six possible outcomes is both even and prime.) The probability of a compound event (A and B) is further discussed in Example 2.

Example 2

Slips of paper, numbered from one to five, are placed in a container and two slips are drawn at random. Determine the probability that both numbers drawn are even if

- the first slip is replaced before the second slip is drawn
- the first slip is not replaced before the second slip is drawn

Solution

- the first slip is replaced before the second slip is drawn

  There are four favorable outcomes {(2, 2), (2, 4), (4, 2), (4, 4)} and $5 \times 5 = 25$ possible outcomes. The probability that both numbers drawn are even is $\frac{4}{25}$.

  This value can be determined by using the compound event theory. Let A represent the event that the first slip drawn is even, A = {2, 4) and $P(A) = \frac{2}{5}$. Let B represent the event that the second slip drawn is even, B = {2,4} and $P(B) = \frac{2}{5}$.

  When event A and event B are independent, P(A and B) = P(A) × P(B).

The two events are said to be **independent**. This means that the probability of the second event is not affected by the occurrence of the first event.

In this example, $P(A) = \frac{2}{5}$ and $P(B) = \frac{2}{5}$, so $P(A \text{ and } B) = \frac{2}{5} \times \frac{2}{5} = \frac{4}{25}$.

- the first slip is not replaced before the second slip is drawn

There are only two favorable outcomes {(2, 4), (4, 2)} and $5 \times 4 = 20$ possible outcomes. The probability that both numbers will be even is $\frac{2}{20} = \frac{1}{10}$. Compound event theory can be applied again, but the assumption that the number on the first slip drawn is even affects the probability that the second value will be even.

Let A represent the event that the first slip drawn is even; then A = {2, 4} and $P(A) = \frac{2}{5}$. If B represents the event that the second slip drawn is even, then B = {4} or B = {2} depending on what happened on the first draw. Regardless, $P(B) = \frac{1}{4}$. These two events are said to be **dependent**. The probability of getting an even number on the second draw is affected by the fact that an even number was obtained on the first draw. Only one of the four slips remaining would be even.

When event A and event B are dependent, P(A and B) = P(A) × P(B|A). "P(B|A)" is read as the "probability of B given A."

When two events are dependent, the sample space of the second event is changed by the occurrence of the first event. In this example, there are five possible outcomes when determining the probability of event A, but only four possible outcomes when determining the probability of event B.

$$P(A) = \frac{2}{5},\ P(B|A) = \frac{1}{4} \text{ and } P(A \text{ and } B) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10}$$

Additional Examples
Mathematics: Principles and Process 12: pp. 416, 418, 420, 422
Holtmath 12: pp. 363, 365, 369, 370

Exercises

1. In a class of 32 students, names were recorded and placed in a bowl. There are 12 girls in the class.

   a. What is the probability of drawing a boy's name?

   b. What is the probability of not drawing a boy's name?

2. A box contains 18 red chips, 12 blue chips and 20 green chips. If you take a chip from the box, what is the probability it will be the following?

   a. a blue chip b. a red chip c. a blue or a red chip

3. A deck of 52 cards is shuffled. After one card is drawn, it is returned to the deck before the next draw. Calculate the probability of the following:

   a. drawing an ace and then a king b. drawing a two and then another two

4. Repeat question 3 given that the first card drawn is not returned to the deck.

5. A box contains 1000 bulbs. If the probability of receiving a defective bulb is 0.015, how many defective bulbs would you expect to find in the box?

BUILDING BLOCKS SOLUTIONS are at the back of this workbook.

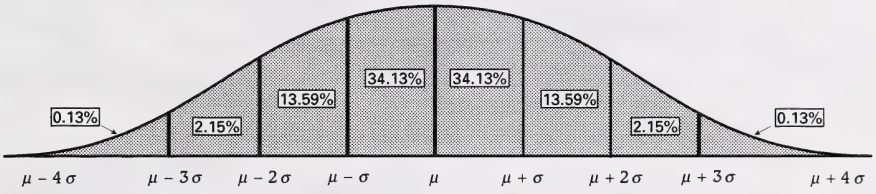
Additional Practice
Mathematics 12: p. 435 (Questions 5, 6, 10)
Mathematics: Principles and Process 12: p. 417 (Question 4)
Holtmath 12: p. 364 (Question 4), p. 367 (Question 7), p. 368 (Question 9), p. 371 (Questions 1, 3, 7)

The Normal Distribution Curve

MATH FACTOR

Watch the Introduction and Video Segment 2a.

The following diagram illustrates how the data in a normal distribution is distributed.



TEXT

Additional Examples
Mathematics 12: p. 464
Mathematics: Principles and Process 12: pp. 449, 450
Holtmath 12: p. 398

Exercises

Answer the following questions with respect to the presented normal distribution curves.

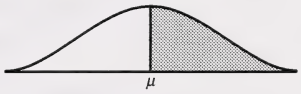
1. a. The mean of the distribution is represented by _____.

   b. The median of the distribution is represented by _____.

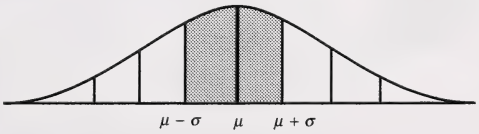
   c. The mode of the distribution is represented by _____.

   d. The total area under the curve represents _____% of the population.

   e. The shaded area represents _____% of the population.



2. a. The standard deviation of the normal distribution represented to the right is represented by _____.



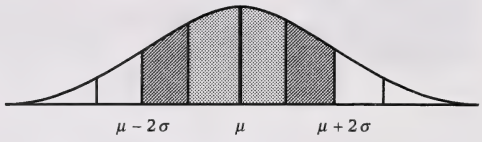
   b. The area under the curve within one standard deviation of the mean represents approximately _____% of the population.

   c. The area under the curve one standard deviation above the mean represents approximately _____% of the population.

d. The area under the curve one standard deviation below the mean represents approximately _____% of the population.

3. a. The area under a normal distribution curve within two standard deviations of the mean is approximately _____% of the population.

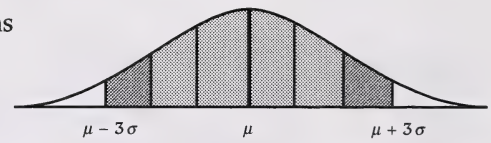
b. The area under the curve two standard deviations above the mean represents approximately _____% of the population.



c. The area under the curve two standard deviations below the mean represents approximately _____% of the population.

d. The area under the curve between one and two standard deviations from the mean represents approximately _____% of the population.

4. a. The area under a normal distribution curve within three standard deviations of the mean represents approximately _____% of the population.



b The area under the curve three standard deviations above the mean represents approximately _____% of the population.

c. The area under the curve three standard deviations below the mean represents approximately _____% of the population.

d. The area under the curve between two and three standard deviations from the mean represents approximately _____% of the population.

EXERCISE 2a SOLUTIONS are at the back of this workbook.

TEXT

Additional Practice
Holtmath 12: p. 399 (Questions 1-4)

Problems Involving Normally Distributed Data

The time that a televised hockey game lasts is normally distributed with an average game time of 2 h 45 min and a standard deviation of 8 min. Determine the probability that a televised hockey game will last at least 181 min.

Watch Video Segment 2b.

Additional Examples
Mathematics 12: p. 465 (Example 2)
Mathematics: Principles and Process 12: p. 450
Holtmath 12: p. 399

Exercises

5. A small electrical company is producing batteries. The length of time that the batteries maintain their charge is normally distributed. If the batteries have an average life of 27 h and a standard deviation of three hours, answer the following:

   a. What is the probability that a battery will maintain its charge longer than 30 h?

   b. If the company plans to replace a battery that does not last at least 18 h, what is the probability that the company will have to replace a battery?

   c. Using the probability calculated in question 5b, and if the company produces 200 000 batteries, how many batteries can the company expect to replace?

6. A small company packages chocolates. The mass of the packages is normally distributed. The mean mass of a package is 45 g with a standard deviation of two grams.

   a. What percent of its production does the company expect to be within two grams of the mean mass?

   b. The company rejects any package when the mass does not fall between 43 g and 49 g. In a production run of 10 000 packages, how many does the company expect to reject?

7. A large company produces electric dishwashers. The life expectancy of the dishwashers is normally distributed with an average life expectancy of ten years with a standard deviation of two years. The company's warranty is based on the fact that it will replace no more than 2.28% of the production. How long, in years, is its warranty?

EXERCISE 2b SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics 12: p. 466 (Questions 7, 8a, 8b, 12a, 12b, 12c, 13)
Holtmath 12: p. 400 (Questions 5-10)

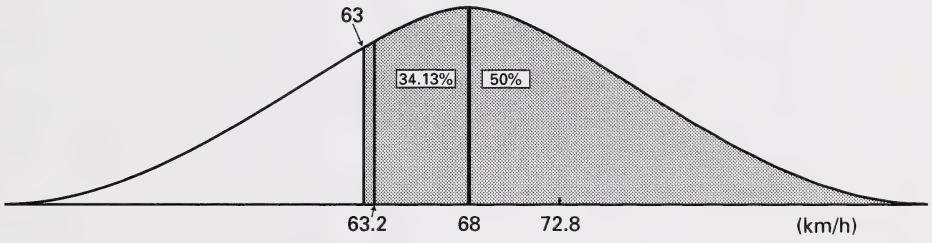
Problems Involving Normally Distributed Data and Estimation

Watch Video Segment 2c.

Example

The speed limit on a road is 60 km/h. The speed of all vehicles on this road is normally distributed with a mean speed of 68 km/h and a standard deviation of 4.8 km/h. If the police give a ticket to all motorists travelling at speeds clocked above 63 km/h, approximately what percentage of motorists will get a ticket?

Solution



The required area below the mean is slightly greater than 34.13%. An estimation of this area would be 35%. Therefore, approximately 35% + 50% = 85% of the motorists will be ticketed.

Note: Estimation is relatively simple when the given value is close to a standard deviation value. In the next video, you will discover how to deal with values for which estimation is more difficult.

Exercises

8. Ball bearings produced by a machine have a mean diameter of 3.50 cm and a standard deviation of 0.10 cm. Ball bearings with diameters between 3.41 cm and 3.59 cm are considered to be acceptable. Assuming a normal distribution of the diameters, estimate the percentage of the ball bearings produced that are acceptable.

WHAT HAVE YOU LEARNED?
Return to the Learner Expectations at the beginning of this lesson and check off the points that you have mastered. Having a firm grasp of the characteristics of a normal distribution is necessary before you continue on to **Lesson 3: Applications of Standard Normal Distribution**.

Solutions

BUILDING BLOCKS

1. a. $\frac{20}{32} = \frac{5}{8}$

   b. $\frac{3}{8}$

2. a. $\frac{6}{25}$

   b. $\frac{9}{25}$

   c. $\frac{3}{5}$

3. a. $\frac{1}{169}$

   b. $\frac{1}{169}$

4. a. $\frac{4}{663}$

   b. $\frac{1}{221}$

5. 15

EXERCISE 2a

1. a. μ

   b. μ

   c. μ

   d. 100

   e. 50

2. a. σ

   b. 68.26

   c. 34.13

   d. 34.13

3. a. 95.44

   b. 47.72

   c. 47.72

   d. 27.18

4. a. 99.74

   b. 49.87

   c. 49.87

   d. 4.30

EXERCISE 2b

5. a. 0.1587 or 15.87% ≐ 16%

   b. 0.0013 or 0.13%

   c. 260

6. a. 68.26%

   b. approximately 1815

7. six years (two standard deviations below the mean)

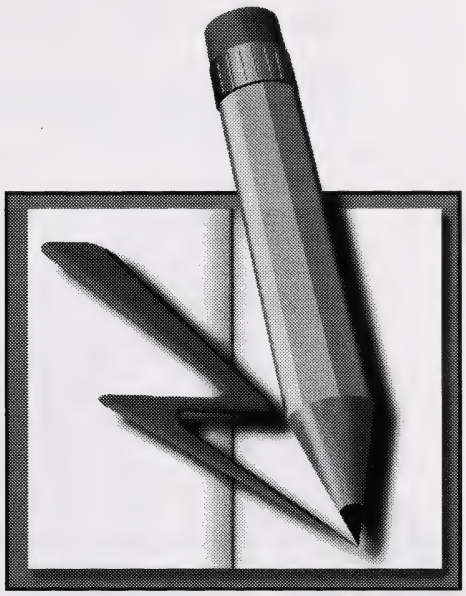
EXERCISE 2c

8. approximately 65%

1997-403 c2 v3

MATH FACTOR

STATISTICS



Lesson 3: Applications of Standard Normal Distribution

Student Workbook

Produced by

THE EDUCATION STATION ACCESS

for

Alberta EDUCATION

This guide accompanies **Applications of Standard Normal Distribution** from the Math Factor series.

Writer: Jerry Polei
Editor: Wilma Gonsalves
Producer: Katherine Rankin

Educators in public educational institutions in Alberta may verbally quote, or use in exams, excerpts from this publication, except material that has been credited to another source. Additional copies of the guide can be obtained, at cost plus handling and postage, from the ACCESS Media Resource Centre.

When ordering videotapes, please use the Basic Program Number (BPN), plus the program number and title(s). In addition, please attach a purchase order to cover duplication fees. The order number and title is

BPN #3689-47
Statistics: Applications of Standard Normal Distribution

For information about other ACCESS materials, the toll-free number for Alberta educators is 1-800-352-8293, or you may write to the following address:

ACCESS
Media Resource Centre
3720 - 76 Avenue
EDMONTON, Alberta T6B 2N9
Telephone: (403) 440-7729
Fax: (403) 440-8899

Inquiries from outside Alberta regarding purchase of this series or other ACCESS materials should be directed to ACCESS Program Sales at the address given above.

© 1996 Alberta Education

ISBN 1-896762-44-1

PRINTED AND BOUND IN CANADA

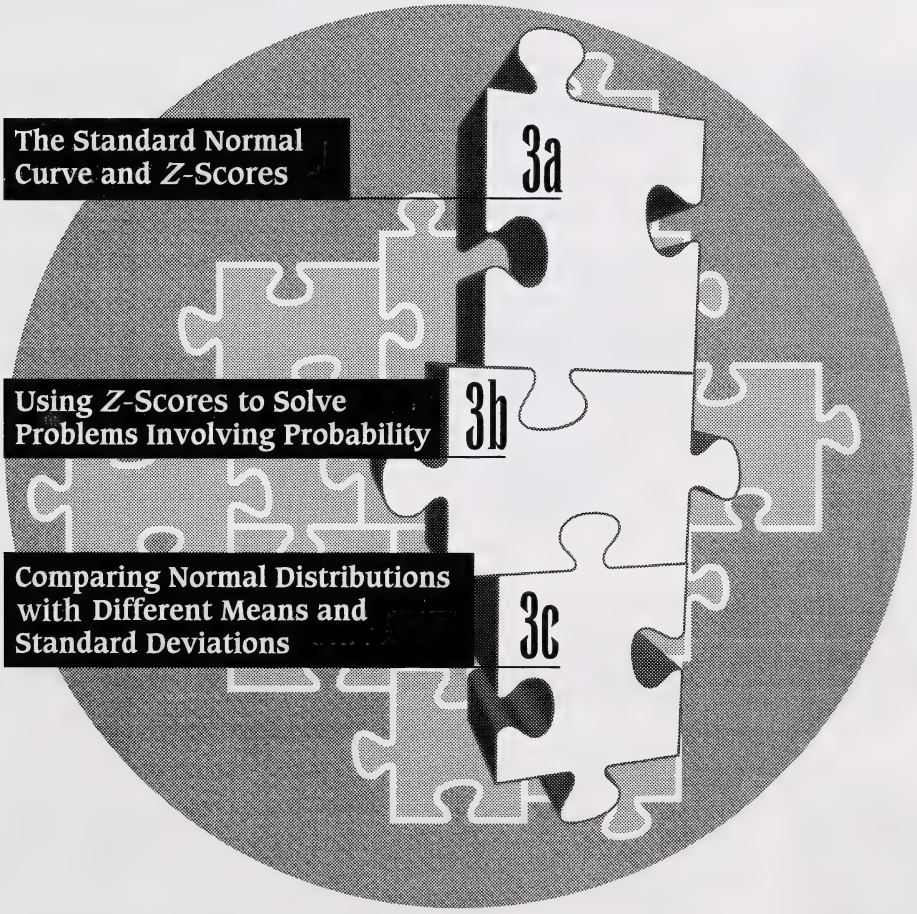
Lesson 3: APPLICATIONS OF STANDARD NORMAL DISTRIBUTION

LEARNER EXPECTATIONS

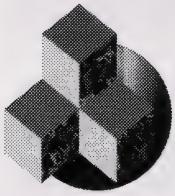
By the end of Lesson 3 you should be able to

- find and apply the standard normal curve and z-scores of data that are normally distributed
- apply z-scores to solve problems involving probability distributions

This lesson has three parts:


The Standard Normal Curve and Z-Scores
3a
Using Z-Scores to Solve Problems Involving Probability
3b
Comparing Normal Distributions with Different Means and Standard Deviations
3c

EXPLANATION OF ICONS



Building Blocks:

Do you remember most of your Mathematics 20? This course builds upon many of those skills and concepts. The Building Blocks icon appears in the workbook to indicate review examples and/or exercises. To refresh your memory, do this section in the workbook before watching the video.

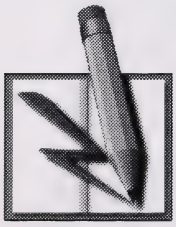
Video:

Whenever this icon appears in the workbook, you should watch the indicated segment of the video program. Preceding this icon, there is often a problem titled "Challenge yourself." Try the problem before watching the video. As you work through a program segment, stop the video to

- challenge yourself to find the solution before it is given
- write down key ideas and examples
- reflect on what is being taught

Challenge Yourself:

This icon sometimes appears on the television screen when a new example is starting. When it appears, it is suggested that you stop the video and try the example yourself before proceeding. For your convenience, sometimes the example is already written in the workbook. The solution will be given in the video.



Student Workbook:

This icon appears on the television screen at the end of teaching segments, when a new skill or sub-skill has been taught. The icon indicates that the workbook contains related examples and/or exercises. Stop the video to study the examples and complete the exercises. The video icon in the workbook will indicate when you should return to the video and move on to learning a new skill.

Excellence:

This icon appears in the workbook to indicate advanced level examples or exercises that you might find challenging. Test your new skills!

Textbook Reference:

This icon appears in the workbook to indicate additional examples and exercises that can be found in the following three textbooks:

- *Mathematics 12 (Addison-Wesley)*
- *Mathematics: Principles and Process 12 (Nelson Canada)*
- *Holtmath 12 (Holt, Rinehart and Winston of Canada, Limited)*

The Standard Normal Curve and Z-Scores

Challenge yourself

Determine the mark on an entrance exam corresponding to a z-score of 1.25 if the marks are normally distributed with $\mu = 70$ and $\sigma = 10$.

MATH FACTOR

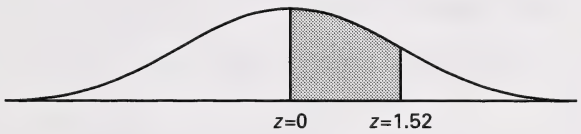
Watch the Introduction and Video Segment 3a.

The standard normal distribution and z-scores are needed to make more accurate approximations of areas. This is especially important when the area is determined by a value or values that are not an exact number of standard deviations from the mean. Also, the standard normal distribution and z-scores allow the comparison of normal distributions that have different means and standard deviations.

It is important to understand the difference between the z-score and the actual area under the curve. Understanding this difference will minimize errors in interpreting problems and in using a z-score table.

Reading Areas from the Table When a *Z*-Score Is Known

The area of shaded portion of the normal distribution pictured can be determined using a z-score table. The area given in a z-score table is usually a four-decimal-place approximation of the area between the mean, $z = 0$, and the z-score to the right of or above the mean. Since the curve is symmetrical about the mean, this value is also the approximate area between the mean and the negative of the z-score being used.



Using the z-score table, the shaded area in this diagram is approximately 0.4357 square units or just 0.4357. Since the total area under the curve is 1, approximately 43.57% of the area is shaded. This can be interpreted to mean that the probability that a piece of data in this population lies between $z = 0$ and $z = 1.52$ is 0.4357. Another way of writing this is $P(0 < z < 1.52) = 0.4357$ or 43.57%.

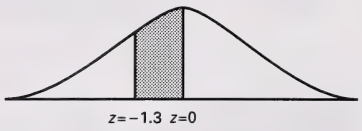
The formula $z = \dfrac{x - \mu}{\sigma}$ can be used to determine any one of z, x, μ or σ provided the other three values are known.

Example 1

Find the probability that a piece of data lies between $z = 0$ and $z = -1.3$.

Solution

Find 1.3 down the left side of the z-score table and record the area as 0.4032.
$\therefore P(-1.3 < z < 0) = 0.4032$



Example 2

For the previous example, if the mean of the distribution under consideration was 6.5 and the standard deviation was 5, what was the data value that generated a z-score of -1.3?

Solution

Substitute the known values into the formula $z = \dfrac{x - \mu}{\sigma}$.

$$-1.3 = \frac{(x - 6.5)}{5}$$
$$-6.5 = x - 6.5$$
$$x = 0$$

The data value was 0.

Additional Examples
Mathematics: Principles and Process 12: p. 452
Holtmath 12: p. 401

Exercises

1. What is the z-score for a data value two standard deviations below the mean?

2. Using a z-score table, determine the following probabilities.

   a. $P(0 < z < 1.5)$

   b. $P(-1.8 < z < 0)$

c. $P(0 < z < 2.3)$

3. Find the area under the standard normal curve between $z = 0$ and $z = -2.5$.

4. A normal distribution has a mean of 85 g and a standard deviation of 20 g. Determine the data value from this distribution that would result in a z-score of 1.75.

EXERCISE 3a SOLUTIONS are at the back of this workbook.

Additional Exercises
Mathematics 12: p. 475 (Questions 5-7, 12-15)
Mathematics: Principles and Process 12: p. 453 (Questions 1-6)
Holtmath 12: p. 401 (Questions 1-6)

Using Z-Scores to Solve Problems Involving Probability

If the length of time that a light bulb continuously operates is normally distributed with a mean of 10 000 hours and a standard deviation of 1100 hours, then determine the probability that a randomly chosen light bulb will operate continuously for at least one year.

Watch Video Segment 3b.

Example 1

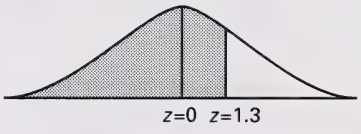
Determine $P(z < 1.3)$.

Solution

The shaded area equals the area below the mean plus the area between the mean and $z = 1.3$.

$$\begin{aligned}\text{shaded area} &= 0.5 + 0.4032 \\ &= 0.9032\end{aligned}$$

Therefore, $P(z < 1.3) = 0.9032$.



Example 2

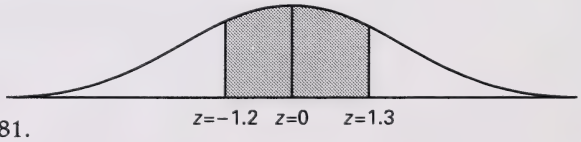
Determine $P(-1.2 < z < 1.3)$.

Solution

The shaded area equals the area between the mean and -1.2 plus the area between the mean and 1.3.

$$\begin{aligned}\text{shaded area} &= 0.3849 + 0.4032 \\ &= 0.7881\end{aligned}$$

Therefore, $P(-1.2 < z < 1.3) = 0.7881$.

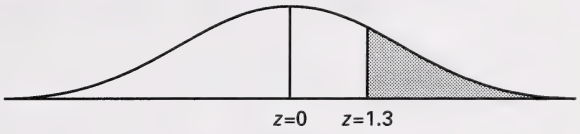


Example 3

Determine $P(z > 1.3)$.

Solution

The shaded area equals the total area above the mean minus the area between the mean and $z = 1.3$.



$$\text{shaded area} = 0.5 - 0.4032$$
$$= 0.0968$$

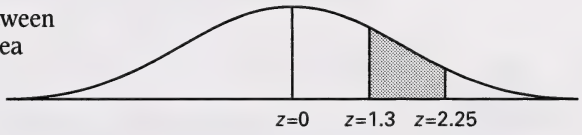
Therefore, $P(z > 1.3) = 0.0968$.

Example 4

Determine $P(1.3 < z < 2.25)$.

Solution

The shaded area equals the area between the mean and $z = 2.25$ minus the area between the mean and $z = 1.3$.



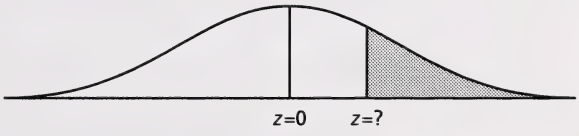
$$\text{shaded area} = 0.4878 - 0.4032$$
$$= 0.0846$$

Therefore, $P(1.3 < z < 2.25) = 0.0846$.

Given an area or a probability, it is often necessary to determine a corresponding z-score from the table. Make sure the area you are using is, in fact, the area between the mean and the desired z-score.

Example 5

The shaded area in the diagram to the right is 0.2850. Determine the z-score that limits this area.



Solution

The shaded area equals the total area above the mean minus the area between the mean and z. So, $0.2850 = 0.5 -$ the area between the mean and z.

The area between the mean and the z-score will be $0.5 - 0.2850 = 0.2150$.

Searching the body of the z-score table for this value indicates that a z-score of approximately 0.57 will determine the shaded area.

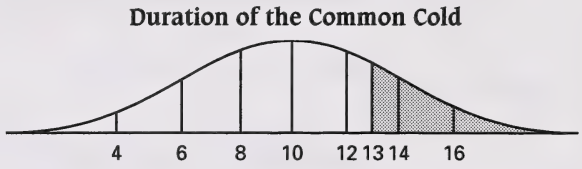
Example 6

The length of time that a common cold lasts is normally distributed. If the mean duration of a common cold is ten days with a standard deviation of two days, then what is the probability that a cold will last longer than thirteen days?

Solution

Sketch a diagram. You need to determine the $P(x > 13)$. Determine the z-score for 13.

$$z = \frac{(13 - 10)}{2}$$
$$= 1.5$$



So, $P(x > 13) = P(z > 1.5)$.

Now determine the area under the standard normal curve for $z > 1.5$.

This is approximately $0.5 - 0.4332 = 0.0668$. $P(x > 13) = P(z > 1.5) = 0.0668$. The probability of having a common cold longer than thirteen days is 0.0668 or almost 7%.

Example 7

A firm manufactures car stereos. The life of the car stereos is normally distributed with a mean life of ten years and a standard deviation of 1.25 years. If the probability that a stereo will last longer than the warranty time is 0.99, then what is the warranty time correct to the nearest half year?

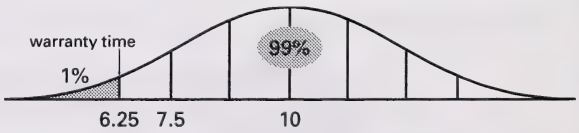
Solution

Sketch a diagram. 99% of the stereos last longer than the warranty time; therefore, 1% last less. The area under the standard normal curve from the mean to the warranty time would be $0.5 - 0.01 = 0.49$.

Determine the z-score for an area of 0.49 from a z-score table. $z \doteq 2.33$. Since the x score is below the mean, the z-score you need is -2.33.

Use the z-score formula to solve for t, the warranty time.

$$-2.33 = \frac{(t - 10)}{1.25}$$
$$-2.9125 = t - 10$$
$$t = 7.0875$$



The shaded area is 0.01 of the total area.

The warranty time is probably seven years.

Additional Examples
Mathematics 12: p. 471 (Example 1), p. 472 (Example 2)
Mathematics: Principles and Process 12: pp. 454, 455
Holtmath 12: pp. 402, 403

Exercises

5. Find the probability of obtaining the following z-scores in a standard normal distribution.

   a. $P(z < 0)$

   b. $P(0 < z < 2)$

   c. $P(z > 1.80)$

   d. $P(-0.89 < z < 1.24)$

   e. $P(-2.37 < z < -0.92)$

6. A firm manufactures batteries with a mean life of 35 h and a standard deviation of 6.5 h. If the life of the batteries is normally distributed, then calculate the probability that a battery is expected to last for the following periods of time.

   a. less than 32 h

   b. less than 43 h

   c. more than 47 h

   d. more than 33 h

7. The expected working years of a refrigerator are normally distributed with a mean of seven years and a standard deviation of one year. In one year, a manufacturer sold 1200 refrigerators.

   a. Of the 1200 refrigerators sold, determine the number of refrigerators expected to work for longer than eight years.

   b. If the manufacturer will replace any refrigerator that stops working before it is five years old, how many refrigerators might be replaced?

8. A manufacturer of thumbtacks sells them in a box. The advertised number of tacks in each box is 100. The number of tacks in a box is normally distributed with a mean of 103 and a standard deviation of 2. Determine the probability that a box will contain 100 or more tacks.

EXERCISE 3b SOLUTIONS are at the back of this workbook.

Additional Exercises
Mathematics 12: p. 475 (Question 15)
Mathematics: Principles and Process 12: pp. 455-457 (Questions 1-12)
Holtmath 12: p. 403 (Questions 1-9)

Comparing Normal Distributions with Different Means and Standard Deviations

Challenge yourself

A student writes an exam along with 451 other people. The scores for the exam are normally distributed with a mean score of 67% and a standard deviation of 12%. How many students scored in the 80's?

MATH FACTOR

Watch Video Segment 3c.

Example 1

If the final marks in both biology and math are normally distributed, then which of the following is the better standing: a mark of 80% in biology, where the mean is 75% and the standard deviation 12%, or a mark of 75% in math, where the mean is 70% and the standard deviation is 15%?

Solution

Find the z-score for each result.

$$z_{\text{bio}} = \frac{5}{12} = 0.42 \qquad z_{\text{math}} = \frac{5}{15} = 0.33$$

Since the z-score for biology is higher than the z-score for math, the standing is better in biology. In this case, a higher z-score indicates a smaller percentage of marks better than the given mark.

Example 2

A standing in the 90th percentile means that 90% of all data lies below this value. Using the information in the previous example, determine the math mark that one would need to have a standing in the 90th percentile.

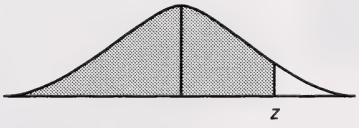
Solution

The area between the mean and the unknown z-score = $0.9 - 0.5 = 0.4$.
This area corresponds to an approximate z-score of 1.28.
Use the z-score formula.

$$1.28 = \frac{x - 75}{15}$$

$$x = 94$$

Therefore, a mark of 94% is required.



Exercises

9. Last year the golf scores of all the golfers on both the front nine and back nine of Lupus Stream Golf Club were found to be normally distributed. There was a mean score of 43 with a standard deviation of two on the front nine, and there was a mean of 41 with a standard deviation of 1.8 on the back nine. Andrea shot a 40 on the front nine and a 38 on the back nine. On which half of the course did Andrea do better relative to the rest of the golfers? (Remember, low scores are desired!)

10. In a competition, the distances that a javelin was thrown were normally distributed with a mean throw of 40.8 m and a standard deviation of 1.0 m. Ken I. Tossett, one of the competitors, threw the javelin 42.0 m. The next day, at the same competition, he participated in the discus throw and threw the discus 38.4 m. If the distances that a discus was thrown were also normally distributed, and the mean distance on the discus throw was 37.6 m and the standard deviation was 1.1 m, then in which competition did Ken have the better standing?

EXERCISE 3c SOLUTIONS are at the back of this workbook.

TEXT

Additional Exercises
Holtmath 12: p. 401 (Questions 5, 6)

WHAT HAVE YOU LEARNED?
Return to the Learner Expectations at the beginning of this lesson and check off the points that you have mastered. Being able to solve problems that involve probability and a normal distribution will promote success as you move on to **Lesson 4: Review**.

Solutions

EXERCISE 3a

1. -2

2. a. 0.4332

   b. 0.4641

   c. 0.4893

3. 0.4938

4. 120 g

EXERCISE 3b

5. a. 0.5

   b. 0.4772

   c. 0.0359

   d. 0.7058

   e. 0.1699

6. a. 0.3228

   b. 0.8907

   c. 0.0322

   d. 0.6217

7. a. $z = 1$ and the area to the right of 1 on the standard normal curve

   $\doteq 0.1587$, $0.1587 \times 1200 \doteq 190$

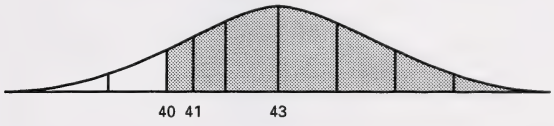
   b. $z = -2$ and the area to the left of -2 on the standard normal curve

   $\doteq 0.0228$, $0.0228 \times 1200 \doteq 27$

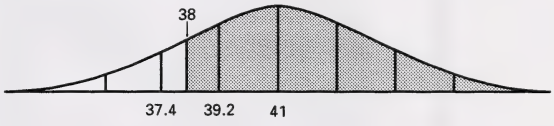
8. 0.9332

EXERCISE 3c

9. $z = -1.5$ and the area under the standard normal curve to the right of -1.5 is 0.9332.



$z = -1.67$ and the area under the standard normal curve to the right of -1.67 is 0.9525.



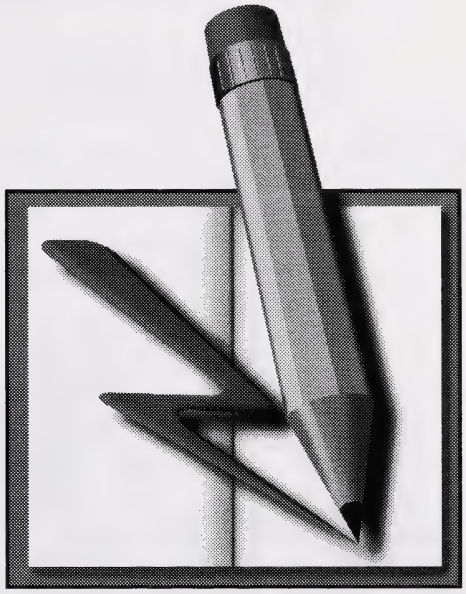
Andrea did better on the back nine. Relative to the rest of the golfers on the back nine, her score was lower than approximately 95.25% of the other golfers.

10. Ken had the better standing in javelin.

L997-403 c2 v4

MATH FACTOR

STATISTICS



Lesson 4: Review

Student Workbook

Produced by

THE EDUCATION STATION ACCESS

for

Alberta EDUCATION

This guide accompanies **Review** from the Math Factor series.

Writer: Jerry Polei
Editor: Wilma Gonsalves
Producer: Katherine Rankin

Educators in public educational institutions in Alberta may verbally quote, or use in exams, excerpts from this publication, except material that has been credited to another source. Additional copies of the guide can be obtained, at cost plus handling and postage, from the ACCESS Media Resource Centre.

When ordering videotapes, please use the Basic Program Number (BPN), plus the program number and title(s). In addition, please attach a purchase order to cover duplication fees. The order number and title is

BPN #3689-48
Statistics: Review

For information about other ACCESS materials, the toll-free number for Alberta educators is 1-800-352-8293, or you may write to the following address:

ACCESS
Media Resource Centre
3720 - 76 Avenue
EDMONTON, Alberta T6B 2N9
Telephone: (403) 440-7729
Fax: (403) 440-8899

Inquiries from outside Alberta regarding purchase of this series or other ACCESS materials should be directed to ACCESS Program Sales at the address given above.

© 1996 Alberta Education

ISBN 1-896762-45-x

PRINTED AND BOUND IN CANADA

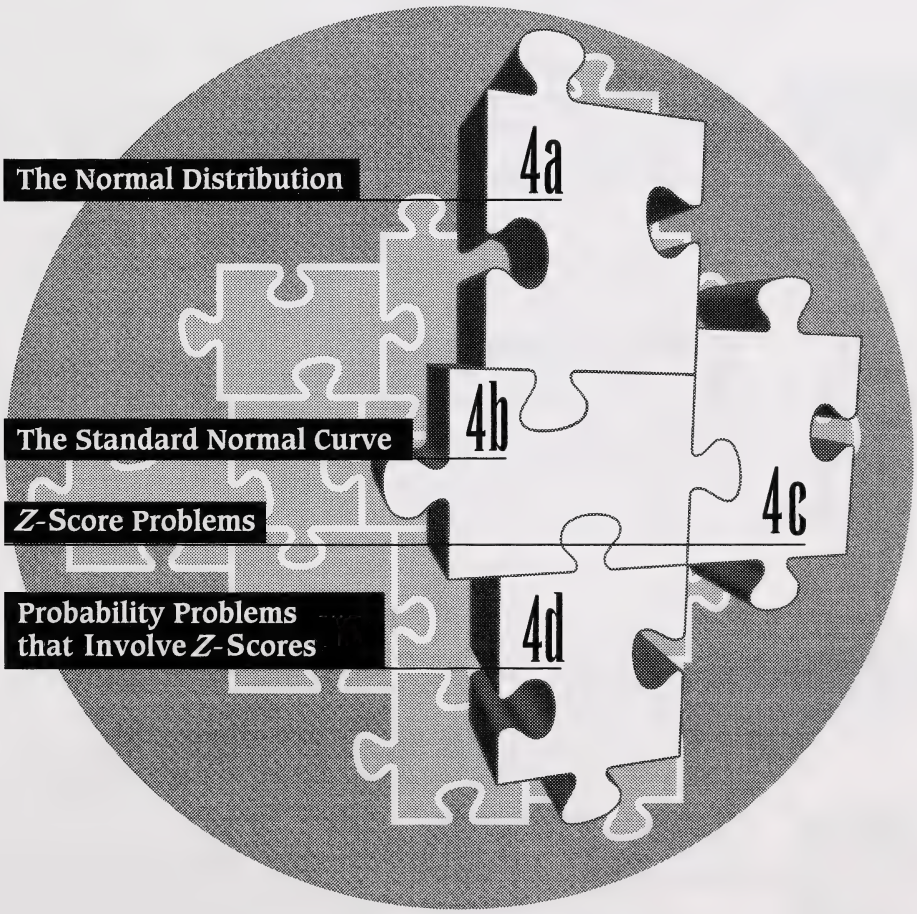
Lesson 4: REVIEW

LEARNER EXPECTATIONS

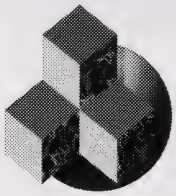
By the end of Lesson 4 you should be able to

- apply newly acquired skills and concepts with regard to the normal curve, z-scores and probability
- use various methods to organize, display and analyze raw data

This lesson has four parts:



EXPLANATION OF ICONS



Building Blocks:

Do you remember most of your Mathematics 20? This course builds upon many of those skills and concepts. The Building Blocks icon appears in the workbook to indicate review examples and/or exercises. To refresh your memory, do this section in the workbook before watching the video.

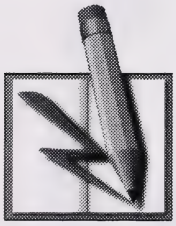
Video:

Whenever this icon appears in the workbook, you should watch the indicated segment of the video program. Preceding this icon, there is often a problem titled "Challenge yourself." Try the problem before watching the video. As you work through a program segment, stop the video to

- challenge yourself to find the solution before it is given
- write down key ideas and examples
- reflect on what is being taught

Challenge Yourself:

This icon sometimes appears on the television screen when a new example is starting. When it appears, it is suggested that you stop the video and try the example yourself before proceeding. For your convenience, sometimes the example is already written in the workbook. The solution will be given in the video.



Student Workbook:

This icon appears on the television screen at the end of teaching segments, when a new skill or sub-skill has been taught. The icon indicates that the workbook contains related examples and/or exercises. Stop the video to study the examples and complete the exercises. The video icon in the workbook will indicate when you should return to the video and move on to learning a new skill.

Excellence:

This icon appears in the workbook to indicate advanced level examples or exercises that you might find challenging. Test your new skills!

Textbook Reference:

This icon appears in the workbook to indicate additional examples and exercises that can be found in the following three textbooks:

- *Mathematics 12 (Addison-Wesley)*
- *Mathematics: Principles and Process 12 (Nelson Canada)*
- *Holtmath 12 (Holt, Rinehart and Winston of Canada, Limited)*

The Normal Distribution

Challenge yourself

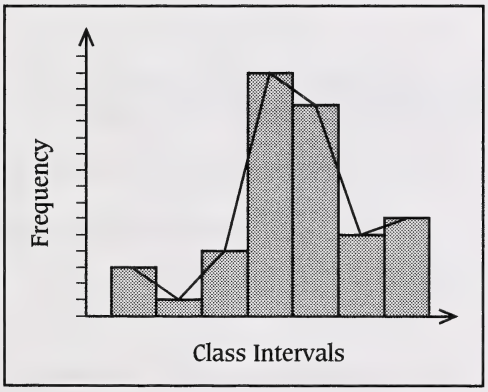
Determine how many of the 950 units sold during a sales promotion will last longer than five years if the mean life expectancy of a unit is 48 months with a standard deviation of 12 months and the life expectancy of the units is normally distributed.

MATH FACTOR

Watch the Introduction and Video Segment 4a.

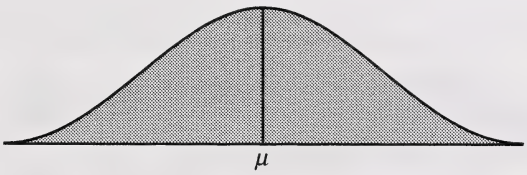
When a histogram or frequency polygon of sample data obtained from normally distributed data is drawn, the resulting diagram often displays some symmetry.

Histogram and Frequency Polygon

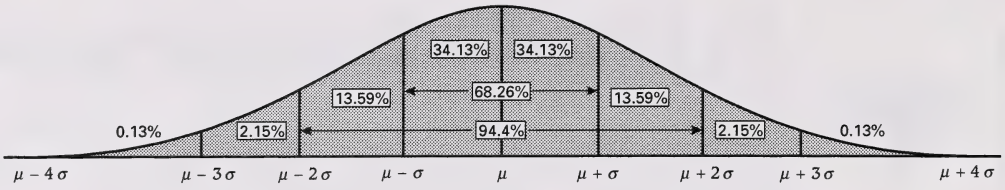


If more samples are taken, the results averaged and smaller intervals used, the graph approaches a bell curve.

Typical Bell Curve for a Normal Distribution



The curve is symmetrical about the mean of the data. The mean, the median and the mode of normally distributed data are equal.



The normal distribution has other properties.

- It has been determined that approximately 68.26% of the data lie under the curve within one standard deviation of the mean. It follows that 34.13% of the data lie one standard deviation above the mean and 34.13% of the data lie one standard deviation below the mean.
- Similarly, 95.44% of the data lie within two standard deviations of the mean. It follows that 13.59% of the data lie within the area between one and two standard deviations above or below the mean.

You will determine and record the remaining properties in question 1 of the following exercise.

Exercises

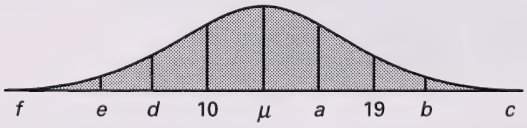
1. Complete the following statements about the normal distribution curve. Use the preceding information given on the normal distribution curve.

   a. Approximately _____% of the data lie within three standard deviations of the mean.
   b. Approximately _____% of the data lie within four standard deviations of the mean.
   c. The total area under the curve can be considered to represent _____% of the data.

2. In a population that is normally distributed, which of the following is true?

   A. Only the mean and the mode, but not the median, are equal.
   B. Only the mean and the median, but not the mode, are equal.
   C. The mean, median and the mode are all different.
   D. The mean, median and the mode are all equal.

3. In the following diagram, each of the vertical lines represents standard deviations in a normal curve. Determine the standard deviation and the values of a, b, c, d, e, f and μ.



4. An examination where the results were normally distributed was scaled up by adding 8% to everyone's mark.

   a. What effect did this have on the examination's mean?

   b. What effect did this have on the examination's standard deviation?

5. An entrance examination with a mean of 65% and a standard deviation of 10% was written by 1000 individuals. Assuming that the results were normally distributed, approximately how many of the individuals who wrote the examination obtained a mark between 55% and 85%?

EXERCISE 4a SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics: Principles and Process 12: p. 470 (Question 1)
Holtmath 12: p. 414 (Question 6)

The Standard Normal Curve

Challenge yourself

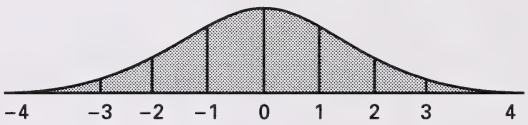
The results of a test were normally distributed and were scaled. Determine a student's new mark given the following:

- The mark the student obtained on the test was 64%.
- The test to be scaled had a mean of 56% and a standard deviation of 6%.
- The marks were to be scaled to a new mean of 65% and a standard deviation of 8%.

MATH FACTOR

Watch Video Segment 4b.

Every normal distribution can be transformed so that its **bell curve** has a **mean of zero** and a **standard deviation of one**. This results in what is referred to as the standard normal curve.



This translation is made using the formula $z = \frac{x - \mu}{\sigma}$. This is the z-score formula used in the videos.

Remember, the z-score tells you how many standard deviations above or below the mean the data value is, depending on whether the z-score is positive or negative. The standard normal curve allows you to use z-score tables or calculators to determine the area under the standard normal curve for any interval.

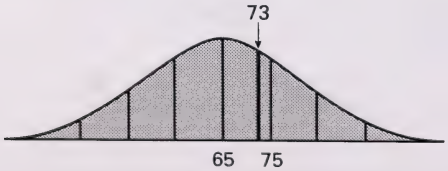
Example 1

The results of a particular exam were normally distributed with a mean of 52% and a standard deviation of 8%. After the exam results were scaled, the mean was 65% and the standard deviation was 10%. One student received a mark of 73% after the exam results were scaled. What was the student's original mark?

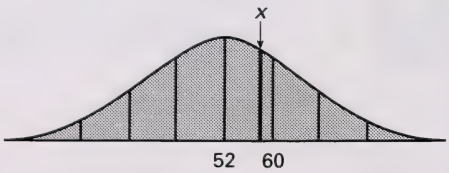
Solution

Sketch a diagram.

After Scaling



Before Scaling



You need to determine x, the student's mark before scaling. Since the student's standing relative to the rest of the students who wrote the exam must stay the same, the student's z-score before the exam was scaled must equal the student's z-score after the exam was scaled.

$$z\text{-score on scaled exam } z = \frac{(73 - 65)}{10} = 0.8 \quad \text{(substituting into } z\text{-score formula)}$$

Use this value to calculate x, the student's original mark before scaling.

$$0.8 = \frac{(x - 52)}{8}$$
$$6.4 = x - 52$$
$$x = 60.4$$

The student's original mark was 60.4%.

Exercises

Assume the data is normally distributed in all of the following cases.

6. Determine the z-score for a mark of 88% on an exam with a mean mark of 65% and a standard deviation of 8%.

7. The z-score for a data value relating to the longevity of a specific brand of automobile tires was 1.56. If the data value was 68 000 km and the standard deviation of the data was 5000 km, what was the average life expectancy of one of these tires?

8. John's scaled mark was 50% on an exam for which his original mark was 42%. The mean mark on this exam before scaling was 48% with a standard deviation of 6%. If the mean mark after scaling was 60%, what was the standard deviation after scaling?

9. The mean length of time that the parts of a microprocessor work is ten years. If the manufacturer gave a seven-year warranty on parts and approximately 2.28% of the processors were returned for repair, what was the standard deviation of the length of time that the processors continued to work?

EXERCISE 4b SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics 12: p. 475 (Question 12)
Holtmath 12: p. 414 (Question 5), p. 416 (Question 26a)

Z-Score Problems

Challenge yourself

If 2800 students wrote an exam where the results were normally distributed with a mean mark of 68% and a standard deviation of 6%, above what mark did the top 100 students place?

MATH FACTOR

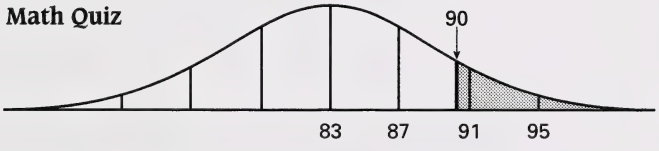
Watch Video Segment 4c.

Example

On a recent math quiz sponsored by the University of Icandoit, the marks were normally distributed with a mean mark of 83% and a standard deviation of 4%. If 12 of the participants achieved a mark above 90%, how many people participated in the math quiz?

Solution

Sketch a diagram.



The shaded area to the right of 90 represents the 12 students who obtained a mark greater than 90%.

Determine the z-score for 90.

$$z = \frac{(90 - 83)}{4}$$
$$= 1.75$$

Determine the area to the right of 90. (The area under the standard normal curve 1.75 standard deviations to the right of the mean is 0.4599 units.)

$$0.5 - 0.4599 = 0.0401$$

So, 0.0401 or 4.01% of the total number of participants is equal to 12.

$$\text{total number of participants} = \frac{12}{0.0401}$$
$$= 299.2$$

Therefore, about 299 people participated in the math quiz.

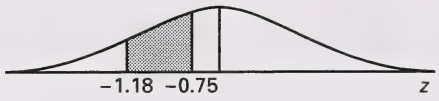
TEXT

Additional Examples
Mathematics: Principles and Process 12: p. 452
Holtmath 12: p. 402 (Questions 1, 2)

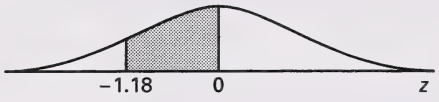
Exercises

10. Determine the area of the shaded portion under the standard normal curve.

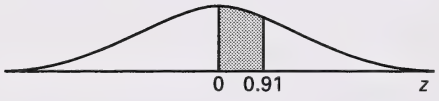
a.



b.



c.



d.

11. If 120 students wrote a test where the results were normally distributed with a mean of 50% and a standard deviation of 12%, how many of the 120 students would be expected to score between 45% and 80%?

12. If the results of a recent math contest sponsored by the University of Waterloo were normally distributed, only 0.02% of the contestants could be expected to obtain a perfect mark of 135. If the standard deviation was 18, what was the mean mark?

EXERCISE 4c SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics: Principles and Process 12: p. 469 (Question 1), p. 470 (Question 2)
Holtmath 12: p. 414 (Question 8)

Probability Problems that Involve Z-Scores

The lives of tires produced by a manufacturer are normally distributed. Determine the probability that a randomly selected tire will qualify for a replacement if the mean tire life is 75 000 km, the standard deviation is 10 000 km and the manufacturer offers free replacement of any of its tires that last less than 60 000 km.

Watch Video Segment 4d.

In the example in 4c, it was pointed out that the area under the curve corresponding to the interval $z > 1.75$ can be interpreted using probability as follows:

> The probability of scoring more than 90% on the math quiz is 0.0401 and this is equal to the probability that the z-score is greater than 1.75.
>
> or $P(x > 90) = P(z > 1.75) = 0.0401$

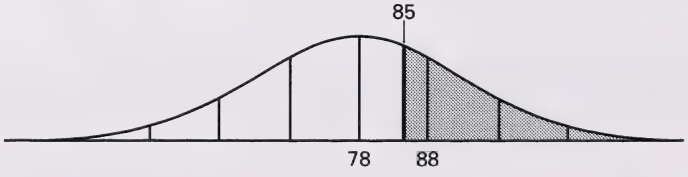
Remember that the sections or intervals under the normal curve represent probabilities as well as areas.

Example

Assuming that the lives of tires are normally distributed, determine the probability that a randomly selected tire will last longer than 85 000 km if the mean tire life is 78 000 km and the standard deviation is 10 000 km.

Solution

Sketch a diagram.



You need to determine $P(x > 85) = P(z > 0.7)$.

The area under the standard normal curve from $z = 0$ to $z = 0.7$ is approximately 0.2580; therefore, the area under the curve to the right of $z = 0.7$ is $0.5 - 0.2580 = 0.242$.

> So, $P(x > 85) = P(z > 0.7) = 0.242$

The probability that a randomly selected tire will last longer than 85 000 km is 0.242.

Additional Examples
Mathematics 12: pp. 471, 472
Mathematics: Principles and Process 12: pp. 454, 455
Holtmath 12: p. 402 (Question 3)

Exercises

13. Determine the following probabilities of a standard normal distribution.

    a. $P(z < -1.02)$

    b. $P(z > -2.04)$

    c. $P(0 < z < 2.45)$

    d. $P(-2.38 < z < -1.89)$

14. The length of time to stand in line to renew a vehicle registration is normally distributed. If the line-up has a mean waiting time of 3.5 min and a standard deviation of 45 s, then what is the probability that you will wait in line for more than 2 min?

15. The "lives" of stationary exercise machines produced by a particular manufacturer are normally distributed. The manufacturer has guaranteed the stationary exercise machines for 20 years.

    a. If the average life of a machine is 23 years with a standard deviation of 1.5 years, what is the probability that a machine selected at random will be returned under conditions of the warranty?

    b. If 1 000 000 machines were manufactured in 1994, how many of these machines can the manufacturer expect to be returned?

EXERCISE 4d SOLUTIONS are at the back of this workbook.

Additional Practice
Mathematics: Principles and Process 12: p. 470 (Question 2)
Holtmath 12: p. 416 (Question 26b)

WHAT HAVE YOU LEARNED?
Return to the Learner Expectations at the beginning of this lesson and check off the points that you have mastered. The Statistics Module is relatively straightforward. It would be wise to practise the concepts and skills you have learned by working through some or all of the statistics-related questions in two or three Mathematics 30 Diploma Examinations.

Solutions

EXERCISE 4a

1. a. 99.74%

   b. 100%

   c. 100%

2. D

3. a. $\sigma = 3$, $\mu = 13$, $a = 16$, $b = 22$, $c = 25$, $d = 7$, $e = 4$, $f = 1$

4. a. The mean would be increased by 8%.

   b. There would be no effect on the standard deviation because the spread of the data has not changed.

5. approximately 819 individuals

EXERCISE 4b

6. $z = \dfrac{(88 - 65)}{8}$

   $\therefore z = 2.88$

7. $1.56 = \dfrac{(68\,000 - \mu)}{5000}$

   $\therefore \mu = 60\,200$ km

8. $z_{\text{unscaled}} = z_{\text{scaled}}$, so $\dfrac{(42 - 48)}{\sigma} = \dfrac{(50 - 60)}{\sigma}$

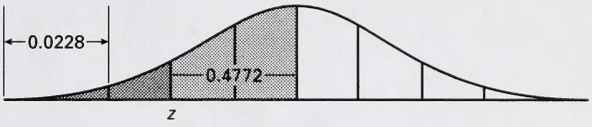
   and $\sigma = 10$. The standard deviation is 10%.

9. Recall that 2.28% of all data in a normal distribution lie below two standard deviations of the mean or determine the z-score that corresponds to the area,
   $(0.5 - 0.0228) = 0.4772$ (below the mean).

   Either way, $z = -2$. Then use the z-score formula.

   $$-2 = \frac{(7 - 10)}{\sigma}$$

   $$\sigma = 1.5$$



The standard deviation is 1.5 years

EXERCISE 4c

10. a. 0.1076 b. 0.3810

    c. 0.3186 d. 0.2380

11. The area under the curve between 45% and 80% is equal to the area under the standard normal curve between $z = -0.42$ and $z = 2.5$. This area is approximately 0.6566 or 65.66% of the total area or population.
    65.66% of 120 = 78.79
    Therefore, approximately 79 students would be expected to score between 45% and 80%.

12. Determine the z-score that corresponds to an area of $(0.5 - 0.0002) = 0.4998$.
    $z = 3.49$
    Then use the z-score formula to obtain $\mu = 72.2$.
    The mean mark was 72.2.

EXERCISE 4d

13. a. 0.1539 b. 0.9793

    c. 0.4929 d. 0.0207

14. $P(x > 2 \text{ min}) = P(z > -2)$
    $= 0.4772 + 0.5$
    $= 0.9772$

15. a. $P(x < 20 \text{ years}) = P(z < -2)$
    $= 0.5 - 0.4772$
    $= 0.0228$

    b. 2.28% of $1\,000\,000 = 0.0228 \times 1\,000\,000$
    $= 22\,800$
    Approximately 22 800 machines may be returned.